



Makale / Research Paper

Derin Öğrenme Yöntemleri Kullanılarak Havadan Elde Edilen Görüntüler Üzerinde Nesne Tespiti

Kemal TÜRKARSLAN^{1a}, Fırat HARDALAC^{2b}

¹ASELSAN A.Ş. Ankara, Türkiye, Gazi Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü. Ankara/TÜRKİYE

²Gazi Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü. Ankara/TÜRKİYE
h.kemal@live.com

Received/Geliş: 24.06.2022

Accepted/Kabul: 18.11.2022

Öz: Son dönemlerde insan hayatı içerisinde gelişmiş uydu sistemlerinin ve insansız hava aracı teknolojilerinin kullanımı günden güne önemli derecede artmaktadır. Bu sistemler üzerinden havadan elde edilen görüntüler savunma sanayii, şehir planlama, tarım, film endüstrisi, eğlence, petrol ve maden arama gibi birçok alanda kullanılmaktadır. Bu gelişmeler doğrultusunda havadan elde edilen görüntüler üzerinde hem nitelik hem nicelik anlamında da artış olmuştur. Bununla beraber günümüzde bilgisayarlı görü dünyasında yapay zekâ algoritmaları üzerinde en çok çalışılan alanlardan biridir. Bilgisayarlı görü ile havadan elde edilen görüntüler üzerinde nesne tespit ve tanıma işlemi oldukça kolay hale gelmektedir. Bu çalışmada tek aşamalı nesne tespit modellerinden YOLOv5 ve SSD algoritmaları kullanılarak Google Earth, GF-2 ve JL-1 uyduları üzerinden toplanan görüntüler üzerinde uçak, gemi, liman ve küçük araç gibi 15 farklı nesnenin tespit çalışmaları yapılmıştır. Google Colab platformu kullanılarak yapılan bu çalışmalarda görüntü bölme (image split) ve veri artırımı (data augmentation) yöntemleri kullanılarak derin öğrenme modelleri üzerindeki etkisi karşılaştırılmıştır. Deneysel çalışmalar sonrasında elde edilen analiz sonuçları grafikler ile paylaşılmıştır.

Anahtar Kelimeler: Uydu Görüntüleri, Hava Fotoğrafı, Nesne Algılama, Evrişimsel Sinir Ağları, Görüntü Bölme.

Object Detection on Aerial Images Using Deep Learning Methods

Abstract: The use of advanced satellite systems and unmanned aerial vehicle technologies has increased in human life day by day in recent times. Aerial images are used in many areas such as defense industry, city planning, agriculture, film industry, entertainment, oil and mineral exploration. After these advancement, there has been an rise in both quality and quantity on the images obtained from the air. Artificial intelligence and deep learning algorithms are one of the most studied and applied fields of study at the same time. Object detection of aerial images will be easy with computer vision. In this study we used dataset images which 15 different class such as plane, ship, port and small vehicles collected by Google Earth, GF-2 and JL-1 satellites. Next, YOLOv5 and SSD algorithms from single-stage object detection models used for object detection. The effects on deep learning models were compared by using image split and data augmentation methods in experiments using the Google Colab platform. Experimental results were analyzed in detail with graphics.

Keywords: Satellite Images, Aerial Photography, Object Detection, Convolutional Neural Networks, Image Splitting.

1. Giriş

Bu makaleye atf yapmak için

Türkarşlan, K., Hardalaç F., “Derin Öğrenme Yöntemleri Kullanılarak Havadan Elede Edilen Görüntüler Üzerinde Nesne Tespiti” El-Cezeri Fen ve Mühendislik Dergisi 2022, 9(4); 1398-1410.

How to cite this article

Türkarşlan, K., Hardalaç F., “Object Dedection on Aerial Images Using Deep Learning Methods” El-Cezeri Journal of Science and Engineering, 2022, 9.(4); 1398-1410.

ORCID: ^a0000-0002-3132-8581, ^b 0000-0003-1358-0756

Bilgisayar bilimciler 1950'den bugüne dijital bir fotoğraf veya videodan tıpkı insanlar gibi bir anlam çıkarabilmesi ile ilgili çalışmalar yapmışlardır. Bilgisayar görüşü ile ilgili ilk deneyler, daire, üçgen gibi basit nesnelere sınıflarına ayırmak için ilk sinir ağlarından bazılarını kullanarak gerçekleştirilmiştir. Teknolojinin 21. yüzyılın başlangıcı ile hızla gelişmesi ve internetin yaygın hale gelerek veri setlerine kolay erişimin sağlanması sayesinde, bilgisayarlı görü alanı ile ilgili çalışmalar hız kazanarak günümüze kadar ulaşmıştır. [1]

İnsansız hava araçları ve uydu sistemleri teknolojileri ilk olarak askeri alanlarda kullanılmaya başlamıştır. Bu sistemlerin sivil amaçlar için kullanılmaya başlanması 1950'li yıllardan sonra gerçekleşmiştir.[2] Siyah beyaz olarak uzaydan çekilen ilk görüntü ABD tarafından gerçekleştirilmiştir. 1974 yılında teknolojinin son imkanları kullanılarak 900 km mesafeden elde edilen görüntü Şekil-1'de gösterilmektedir.[3]



Şekil 1. Uzaydan çekilen ilk görüntü

Havada bulunan bir nesne üzerinden fotoğraf çekilmesi hava fotoğrafı olarak adlandırılmaktadır. İlk havadan elde edilen görüntü 13 Ekim 1860 yılında 609 metre yükseklikten James Wallace Black ve Samuel Archer King tarafından Boston'da çekilmiştir. "Boston, as the Eagle and the Wild Goose See It" ismini verdikleri bu fotoğraf tarihte ilk havadan elde edilen görüntü olarak yerini almıştır.[4]



Şekil 2. Boston, as the Eagle and the Wild Goose See It isimli fotoğraf karesi[4]

Son yarım asır içerisinde dünyamız, dijitalleşme yolunda hızla yol kat etmiştir. Teknolojinin gelişimi ve yapay zekâ özelinde derin öğrenme algoritmalarının önem kazanması ile bilgisayarla

görme dünyasında daha efektif ve verimli tasarımların geliştirilmesi ihtiyacı ortaya çıkmıştır. Bunun yanı sıra insansız hava araçlarının ve gelişmiş uydu sistemlerinin kullanımı oldukça yaygın hale gelmiştir. Bu sistemler kullanılarak yüksek çözünürlüklü fotoğraf ya da video görüntüleri elde edilmiştir. Bu görüntüler sınır güvenliği, keşif gözetleme gibi askeri alanlarda sıklıkla kullanılmaya başlanmıştır. Askeri amaçların yanı sıra tarımsal uygulamalar, trafikte yol durumu, afet yönetimi, meteoroloji gibi sivil alanlar üzerinde de bilgisayarlı görü kullanılarak insan yaşamının kolaylaştırılması amaçlanmıştır. Yaşanan bu gelişmeler sonrasında derin öğrenme algoritmalarının havadan elde edilen görüntüler üzerinde anlam kazandırılmasına yönelik çalışmalarda büyük ölçüde artış gözlenmiştir.



Şekil 3. Ukrayna sınırının güneyinde bulunan Rusya'nın Primorsko Akhtarsk Hava Üssü'nde konuşlandırılan Su-34 savaş uçakları. [5]

Yapılan çalışmalar sonrasında yerden çekilen fotoğraf veya video görüntüsü üzerinde derin öğrenme tekniklerinin nesne tespiti konusunda oldukça başarılı olduğu gözlenmektedir. Havadan elde edilen görüntüler üzerinde de bu başarımın sağlanması konusunda da çalışmalar sürmektedir. Bu başarımın yakalanması halinde birçok alanda insan hayatına önemli ölçüde katkı sağlayacağı düşünülmektedir. Örnek verilecek olursa milli sınır güvenliği amacıyla gözetleme işi yapan uzman askeri bir personelin masa başında görüntüleri tek tek incelemesi yerine derin öğrenme teknikleri kullanılarak etkili görme sistemlerinin geliştirilmesi sonrasında uzman askeri bir personelin yapacağı işi daha hızlı ve daha doğru sonuçlar üreterek sonuçlandırması olağan hale gelecektir. Bu da hem nitelikli insan gücü hem de zamandan kazanç olacaktır.

2. Literatür Özeti

Bu çalışmaya başlamadan önce konu ile ilgili yapılan literatür taramasında derin öğrenme teknikleriyle nesne tanıma çalışmalarının tek ve iki aşamalı nesne tespit modelleri şeklinde iki ayrı kategoriye ayrıldığı görülmüştür. Tek aşamalı nesne tespiti için sıklıkla kullanılan modellere örnek verildiğinde YOLO (You Only Look Once), SSD(Single Shot Dedector), RetinaNet ve EfficientNet modelleri yaygın olarak kullanılırken, iki aşamalı nesne tespit çalışmalarında RCNN, Fast R-CNN ve Faster R-CNN modelleri en etkin ve zaman açısından hızlı sonuçları vermektedir.

Jinwang Wang ve arkadaşları insansız hava aracından elde edilen görüntüler üzerinden çevre kirliliğine engel olmak için bir yazılım geliştirmişlerdir. Bu yazılım ile doğada bulunan plastik

şişelerin konumlarını tespit etmişlerdir. Kendi oluşturdukları veri seti görüntülerini Faster R-CNN, SSD, YOLOv2 ve RRPN (Rotation Region Proposal Networks) gibi nesne tespit algoritmalarında eğitim ve doğrulamaya tabii tutmuşlardır. Deney sonrasında RRPN algoritmasının en yüksek konumlandırma doğruluğu (%88.6) verdiğini gözlemlemişlerdir.[6]

Suat Toraman çalışmasında havadan elde edilen fotoğraflar ile yaya ve bisikletli sınıflarının tespitine yönelik bir yazılım geliştirmiştir. İlk olarak derin öğrenme ile öznelik çıkarımı yapmıştır. Veri setini daha verimli hale getirmek için görüntü büyüme yöntemi kullanmıştır. Bu yöntem ile İHA'dan elde edilen yaya, bisikletli gibi görüntüler yeniden boyutlandırılarak eğitim gerçekleştirilmiştir. CNN modellerinden AlexNet'te %99.84 doğruluk elde edilirken, VGG16'da %99.78 oranında sınıflandırma doğruluğu elde edilmiştir. [7]

Ersin KILIÇ ve arkadaşları CARPK veri seti üzerinde nokta etiketlemeyle yeni bir veri seti oluşturmuşlardır. Bu yeni veri seti ile insansız hava aracı vasıtasıyla elde edilen fotoğraflardan araç sayımı yapabilen derin öğrenme modeli önermişlerdir. Yapılan çalışmanın başarımının değerlendirilmesi adına oluşturulan yeni veri seti ile deneysel sonuçlar elde edilmiştir. Deneysel çalışmalar sonrasında çerçeve etiketleme yöntemi kullanılarak eğitilen modeller ile önerdikleri yöntemin yaklaşık aynı düzeyde başarımlar elde ettiğini değerlendirmişlerdir. [8]

Erdem Bayhan ve arkadaşları çalışmalarında, insansız hava aracı yardımıyla elde edilen fotoğraflar üzerinde Faster-RCNN ve YOLOv4 mimarisini kullanmışlardır. Yaptıkları çalışma ile askeri açıdan olumsuz değerlendirilen tehditlerin İHA tarafından derin öğrenme tabanlı tespit ve tanıma yöntemleri konusunda deneyler yapmışlardır. Deneysel sonuçlar sonrasında Faster-RCNN algoritmasının nesne tespitinde %93 doğruluk oranı elde ettiği, YOLOv4 algoritmasının ise %88 doğruluk oranı elde ettiği gözlenmiştir.[9]

Hüseyin Seçkin DIKBAYIR ve arkadaşları yaptıkları çalışmalarında, gerçek zamanlı olarak kullanılacak araç tespit yazılımı geliştirmişlerdir. Çalışmada evrimsel sinir ağları yapısı kullanılarak, gerçek zamanlı nesne tespit modellerinden YOLO algoritmasını kullanmışlardır. YOLO algoritmasının performansını arttırmak için evrimsel sinir ağlarından faydalanarak bazı metotlar uygulamışlardır. Testler sonrasında tespit hızını artarken doğruluk oranlarının düştüğü gözlenmiştir.[10]

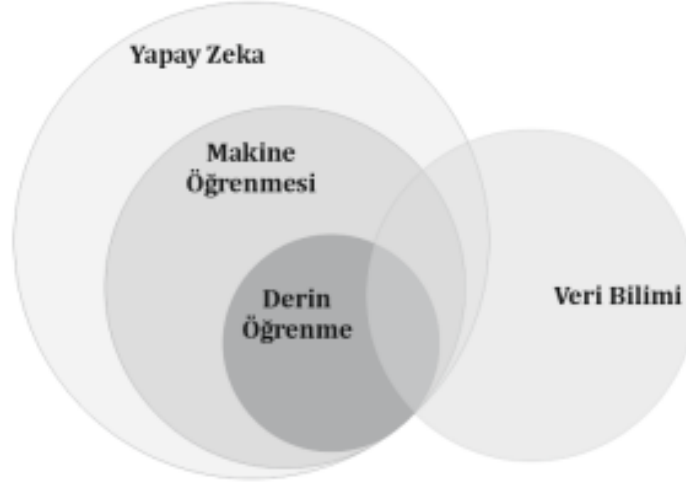
Uğur Talaş ve arkadaşları çalışmalarında nesne tanıma işlemi yapan bir robot geliştirmişlerdir. Uygulamalarında Raspberry Pi3/B mini bilgisayar ile Python dilinde hazırladıkları kodları kullanmışlardır. Nesne tespitinde SSD Mobilenet V1 algoritmasını kullanmışlardır. Testler sonrasında nesne tanıma işlemlerinin başarıyla gerçekleştirildiği görülmüştür. Ancak ortamda bulunan ışığın şiddetine ve giriş görüntü boyutlarına bağlı olarak nesne tanıma işleminin 3-4 saniye gecikme ile gerçekleştiği gözlemlenmiştir. SSD Mobilenet V1 algoritmasının insan, bitki, laptop gibi farklı sınıftaki nesnelere başarı ile tespit ettiği değerlendirilmiştir. [11]

3. Materyal ve Metot

3.1. Derin Öğrenme

Derin öğrenme, yapay zekanın alt dalı olarak kabul edilen bir makine öğrenmesi sınıfıdır. Yapay zeka ile arasında bulunan hiyerarşik yapı Şekil 4'te gösterilmiştir. Derin öğrenmede öznelik çıkarma veya dönüştürme amacıyla birçok yapay sinir ağı katmanı kullanılmaktadır.[12] Bu katmanlarda oluşturulan matematiksel bloklarda oldukça fazla veri işlenmektedir. Derin öğrenme katmanlarında daha önce yapılan pek çok işlem bir arada yürütülerek veri bir sonraki aşama için

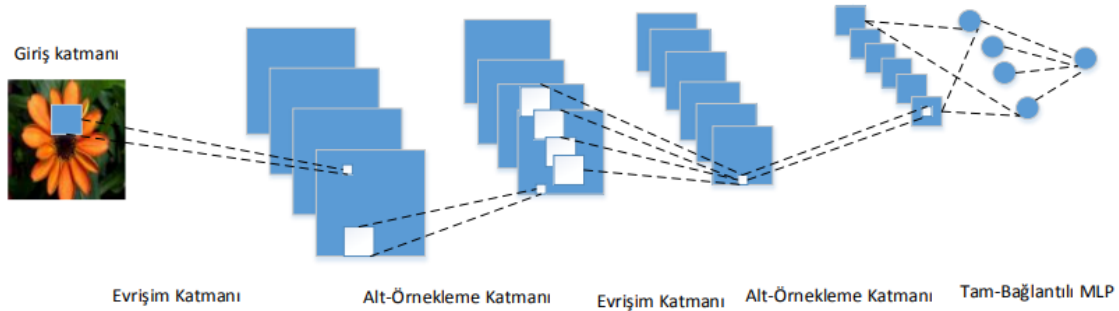
hazır hale getirilmektedir. Bu katmanların sayısının artması işlem karmaşasının artmasına neden olmaktadır. Bu işlem karmaşasının üstesinden gelebilmek amacıyla birçok derin öğrenme mimarisi geliştirilmiştir. En sık kullanılan derin öğrenme mimarileri ise Evrişimsel Sinir Ağları, Uzun Kısa Vadeli Hafıza Ağları ve Tekrarlayan Sinir Ağlarıdır. Bu çalışmada Evrişimsel Sinir Ağları mimarisi üzerinde durulacaktır.



Şekil 4. Yapay zekâ ve alt dalları arasındaki hiyerarşisi gösterimi[13]

3.1.1. Evrişimsel Sinir Ağları

Evrişimsel Sinir Ağları çok katmanlı algılayıcılar kullanan bir yapay sinir ağı türüdür. Evrişimsel Sinir Ağları birden fazla özelliği bulunan bir görüntüyü girdi olarak alıp, görüntüdeki çeşitli özellikleri birbirinden ayırt etmede etkili ve verimli sonuçlar vermesi sebebiyle nesne tanıma, sınıflandırma görevlerinde sıklıkla kullanılan bir derin öğrenme algoritmasıdır. Diğer sinir ağlarından farklı olarak evrişim, özellik çıkarma ve sınıflandırma katmanlarını içermektedir. Şekil 5'te klasik bir evrişimsel sinir ağı yapısı gösterilmiştir. Evrişimsel Sinir Ağları 3 katmandan oluşmaktadır. Bunlar; evrişim, havuzlama ve tam bağlantılı katmanlarıdır.[14]

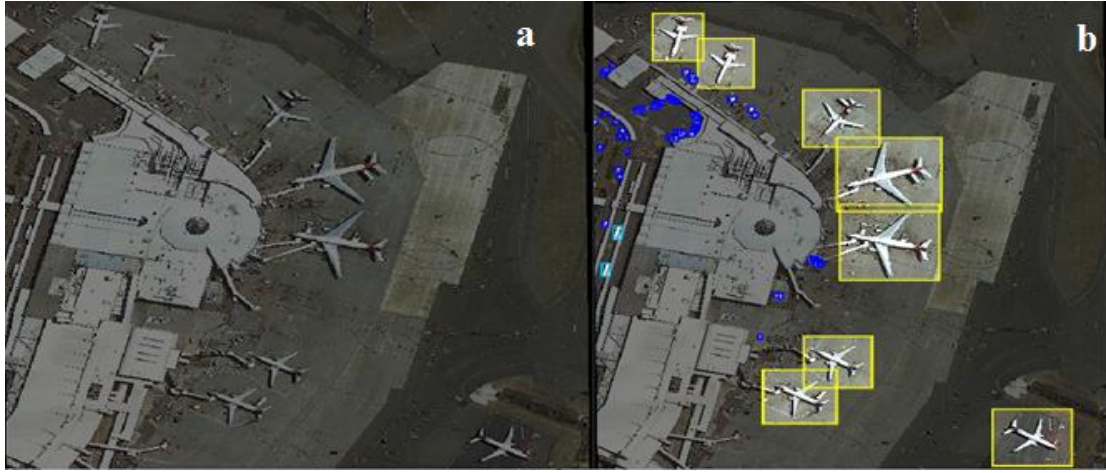


Şekil 5. Evrişimsel Sinir Ağları Yapısı

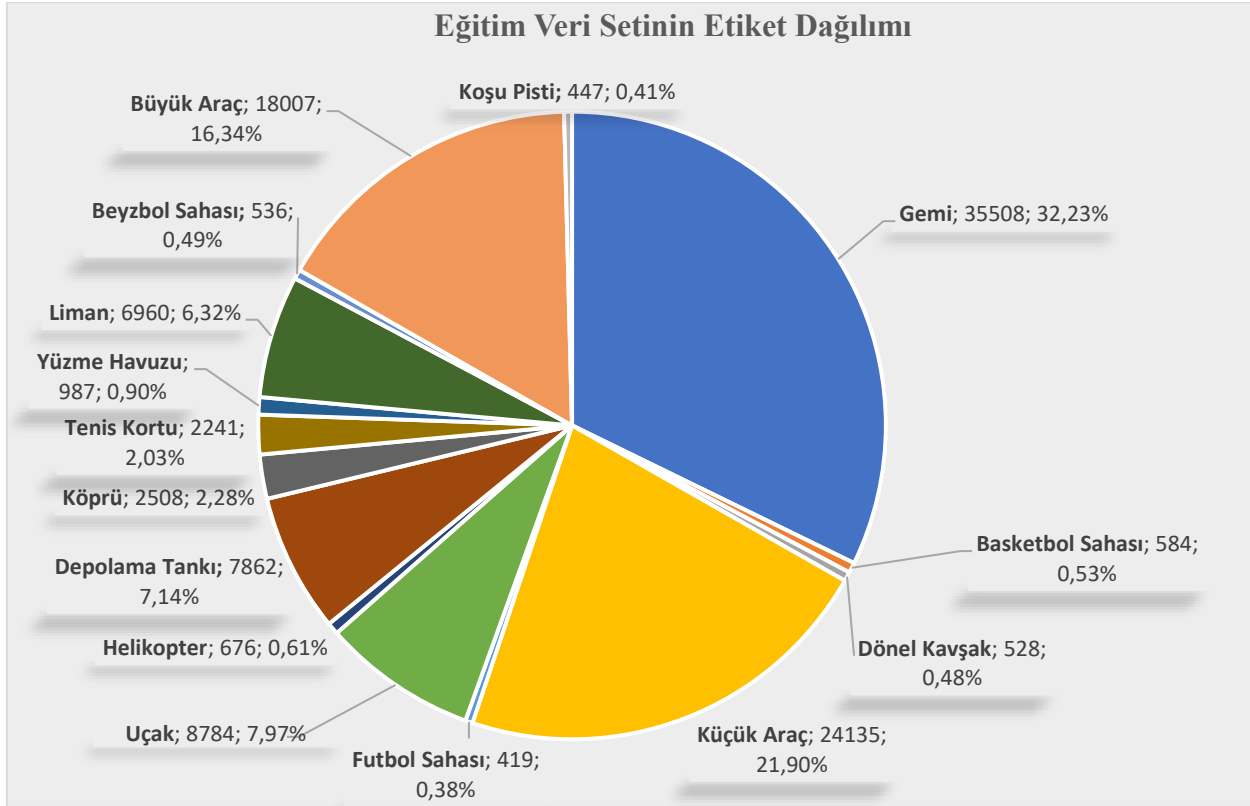
Evrişimsel Sinir Ağı çalışmalarında katmanların çıktı üretmesi için giriş görüntüsünün hepsinin aynı zamanda kullanılmasına gerek yoktur. Evrişimsel Sinir Ağı yapısında bulunan katmanlar içerisinde farklı görevlerde kullanılmak üzere oluşturulan filtreler sayesinde her katmanda evrişim işlemi yapılmaktadır. Bu sayede aynı anda işleme tabi tutulacak öznelik sayısı büyük oranda azaltılabilmektedir. Bu özelliği neticesinde bilgisayarlı görü uygulamalarında diğer derin öğrenme algoritmalarından daha etkili ve verimli sonuçlar ortaya koyabilmektedir [14].

3.2 Veri Seti

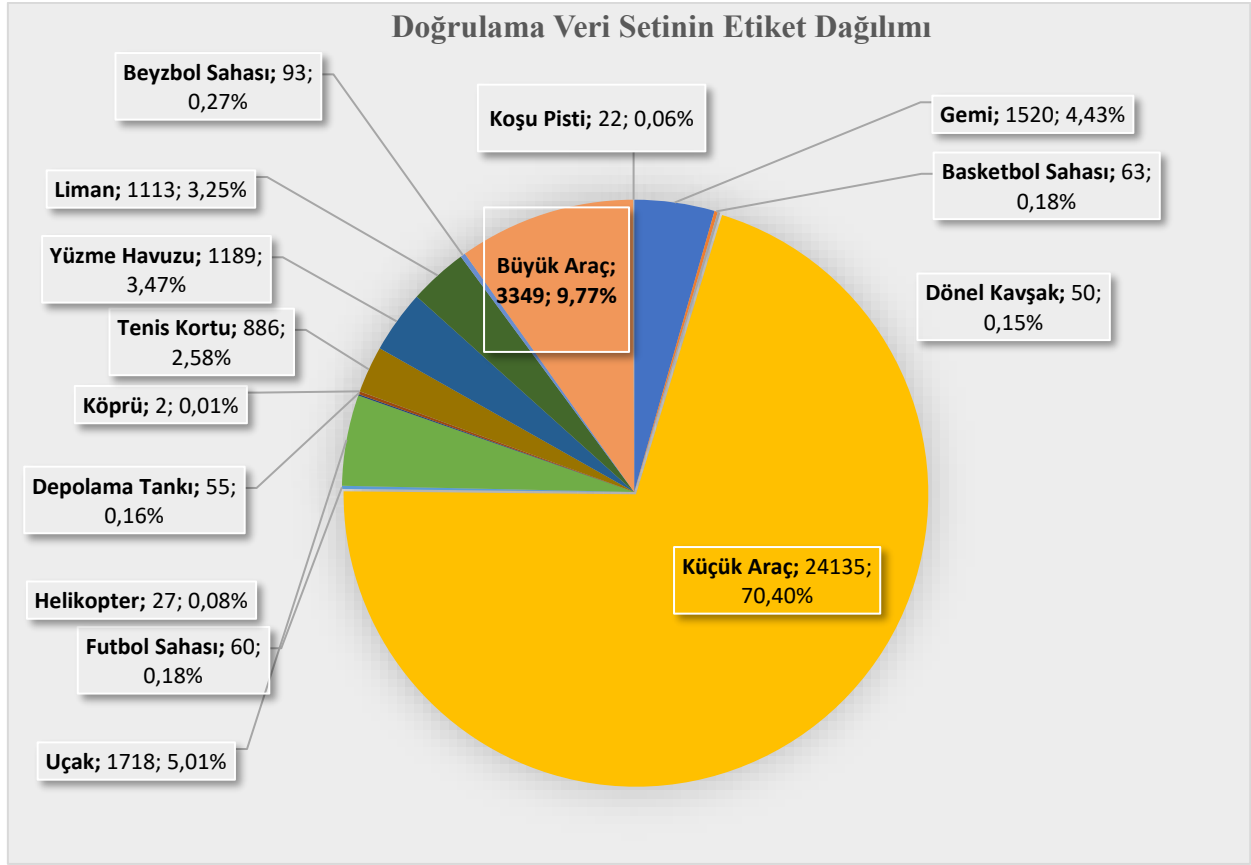
Literatür araştırması sonrası nesne tanıma ve tespiti çalışmalarında genellikle Pascal VOC, MS COCO, CIFAR, ImageNet ve MNIST gibi veri setleri kullanılarak derin öğrenme modellerinin eğitim ve doğrulamaya tabii tutulduğu görülmüştür. Bu deneysel çalışmada ise Çin Kaynaklar Merkezi Uydu Verileri ve Uygulamaları ve Cyclo Media şirketi tarafından Google Earth, GF-2 ve JL-1 uyduları üzerinden havadan elde edilen görüntüler ile oluşturulan DOTA veri seti kullanılmıştır. 1462 adet eğitim, 403 adet doğrulama görüntüsünden oluşan bu veri seti 15 farklı sınıftan oluşmaktadır. DOTA veri seti içerisinde alınan etiketli ve etiketsiz örnek görüntü Şekil 6'da, veri setinin eğitim ve doğrulama setleri içerisinde etiket dağılımları Şekil 7 ve Şekil 8'de gösterilmektedir.



Şekil 6. Etiketsiz (a) ve Etiketli (b) Görüntü [15]



Şekil 7. DOTA Eğitim Görüntülerinin Etiket Dağılımı [15]



Şekil 8. DOTA Doğrulama Görüntülerinin Etiket Dağılımı [15]

3.3 Yöntem

Evrişimsel sinir ağı modelleri genellikle yüksek hız ve verimlilikte çalışan donanımlar gerektirmektedir. Bu sebeple yapılan deneysel çalışmalarda modellerin eğitimi ve doğrulamasında Google tarafından kullanıcılara sunulan bir alt yapı kullanılmıştır. Google Colab Pro adı verilen bu platform tamamen bulut tabanlı olması, yüksek derecede ekran kartı ve RAM kullanımı sunması sebebiyle nesne tespit çalışmaları için oldukça elverişli durumdadır. Colab Pro platformu özellikleri Tablo 1’de gösterilmiştir. Python dilinde hazırlanan kodlar Google Drive üzerinden Colab’a yüklenerek deneyler gerçekleştirilmiştir.

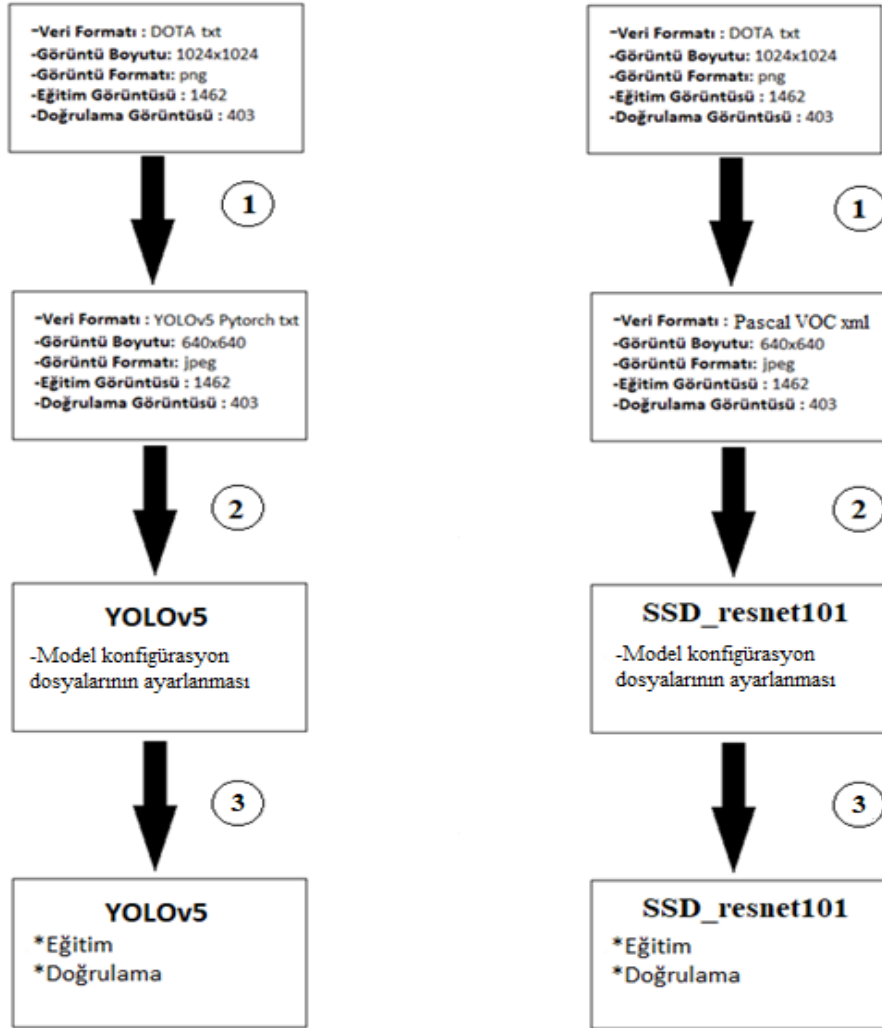
Tablo 1. Google Colab Pro Donanım ve Yazılım Bilgileri

İşletim Sistemi	Linux-5.4.144+-x86_64-with-Ubuntu-18.04-bionic
Sistem Donanımı	Tesla P100-PCIE-16GB
Python Versiyonu	3.7.12

Son zamanlar içerisinde yapılan çalışmalar incelendiğinde nesne tespiti amacıyla kullanılan görüntü işleme algoritmalarında tek aşamalı nesne tespit modelleri arasında YOLO ve SSD algoritmalarının sıklıkla kullanıldığı görülmektedir.[16] SSD algoritması öznitelik haritalarını çıkarma ve daha sonrasında konvolüsyon filtrelerini tespit edilen objelere uygulayarak küçük nesnelere tespit eden bir algoritmadır. İki aşamalı nesne tespit modellerinden Faster R-CNN algoritmasına göre tespit başarısının daha az olmasıyla beraber daha hızlı sonuçlar üretmektedir. SSD algoritması tek aşamalı modeller ile kıyaslandığında YOLO algoritmasından daha başarılı sonuçlar ürettiği görülmektedir. YOLO algoritması girdi görüntüsünü tamamen konvolüsyonel olarak düşünerek, işlem hızını arttırmayı hedeflemektedir. Bu sebeple gerçek zamanlı uygulamalar için kullanımı uygun hale gelmektedir. Resmi direk konvolüsyonel olarak ele alması sebebiyle küçük nesne tespitlerinde

başarımı düşmektedir. [10] COCO veri seti üzerinde yapılan çalışmalar neticesinde YOLO modelinin büyük objelerde %98 oranında başarımlar yakaladığı, küçük objelerde ise %60'lara kadar başarımın düştüğü ifade edilmiştir. [17]

Tek aşamalı nesne tespit modellerinden YOLO ve SSD algoritmalarının nesne tespitinde havadan elde edilen görüntüler üzerindeki etkisini gözlemlemek adına YOLOv5 ve SSD_resnet101 algoritmaları deneysel çalışmalarımızda kullanılmıştır. Özellikle YOLOv5 algoritması ile yapılan çalışmaların literatüre katkısını arttırmak adına bu algoritma ile birçok kombinasyon denenmiş ve deneyler gerçekleştirilmiştir.



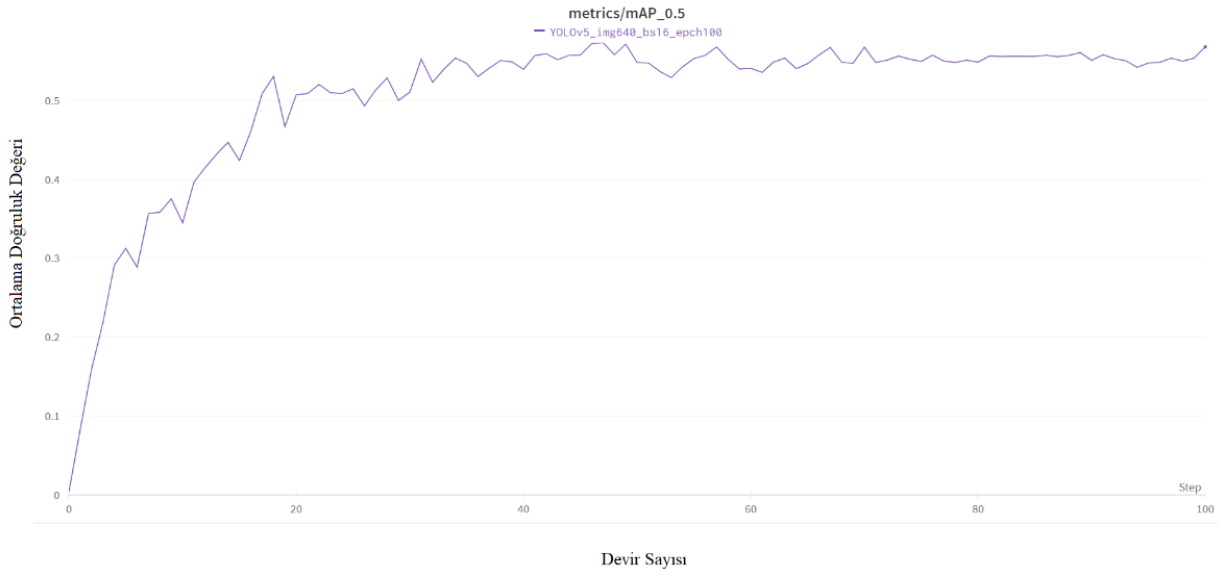
Şekil 9. YOLOv5 ve SSD_resnet101 algoritmalarının DOTA veri seti üzerinde nesne tanıma süreci

İlk olarak DOTA veri setini incelediğimizde veri setinin PNG (Portable Network Graphic) görüntü formatında ve 800×800 ile 20.000×20.000 piksel arasında değişen görüntü boyutlarına sahip olduğu görülmüştür. [15] Şekil 9'da YOLOv5 ve SSD_resnet101 algoritmalarının DOTA veri seti üzerinde nesne tanıma süreci adım adım anlatılmıştır. 1 numara ile gösterilen adımda DOTA veri setinde bulunan 1865 görüntü %70 eğitim, %30 doğrulama seti olacak şekilde bölünmüştür. Evrişimsel sinir ağları modelleri giriş görüntüsü olarak farklı giriş formatlarında çalışmaktadırlar. YOLOv5 algoritması, giriş görüntü bilgilerini içeren dosya formatını .txt olarak belirlerken, SSD algoritması, .xml uzantılı dosya formatı ile çalışmaktadır. Bu sebeple veri seti içerisinde bulunan görüntülerin formatları kullanılacak olan nesne tespit algoritmalarına uygun giriş formatlarına dönüştürülmüştür. Görüntü boyutları 640×640 şeklinde ayarlanarak çalışmalara bu şekilde devam

edilmiştir. 2 numara ile gösterilen adımda ise algoritma modellerinin konfigürasyon ayarları yapılmıştır. Son aşamada DOTA veri seti, her iki model üzerinde eğitime tabii tutularak eğitim sonunda doğrulama veri seti ile doğrulama adımı gerçekleştirilmiştir.

4. Bulgular ve Tartışma

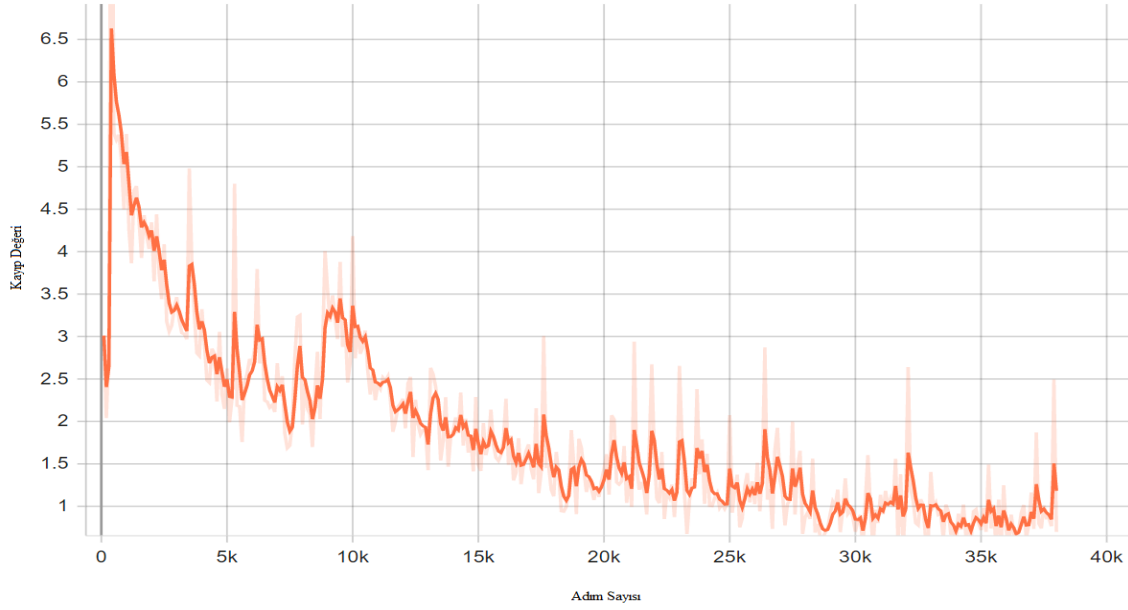
DOTA veri seti kullanılarak ilk olarak YOLOv5 Pytorch modeli ile deneyler gerçekleştirilmeye başlanmıştır. Eğitim ve doğrulama işlemlerinin gerçekleştirileceği Google Colab Pro platformunun kullandığı ekran kartı ve RAM kapasitesi göz önüne alınarak en efektif sonuçların elde edilmesi amacıyla yığın boyutu 16, epok değeri 100 olacak şekilde hiperparametreler ayarlanarak eğitim aşaması başlatılmıştır. Eğitim yaklaşık olarak 7 saat 15 dakikada tamamlanmıştır. Eğitilen modelin doğrulama veri seti üzerindeki sonuçları değerlendirildiğinde, ortalama doğruluk oranı başarısı 0.567 belirlenmiştir. Bu sonuçlar doğrultusunda model, doğrulama veri setinde içerisinde yer alan görüntüler üzerindeki objeleri yaklaşık %56 oranında tespit edip sınıflandırabilmiştir. Şekil 10'da veri setinin eğitimi sonucunda elde edilen ortalama doğruluk oranı grafiği gösterilmiştir.



Şekil 10. YOLOv5 algoritmasının 100 epok eğitim sonrası mAP grafiği

İkinci olarak DOTA veri seti kullanılarak SSD_resnet101 modeli ile deneylere devam edilmiştir. Google Colab Pro platformunun kullandığı ekran kartı ve RAM kapasitesi göz önüne alınarak en efektif sonuçların elde edilmesi amacıyla yığın boyutu 4, adım sayısı 35000 olacak şekilde hiperparametreler ayarlanarak eğitim aşaması başlatılmıştır. Eğitim 4 saat 32 dakika tamamlanmıştır. Eğitilen modelin doğrulama veri seti üzerindeki sonuçları değerlendirildiğinde, ortalama doğruluk oranı başarısı 0.255 belirlenmiştir. Bu sonuçlar doğrultusunda model, doğrulama veri setinde içerisinde yer alan görüntüler üzerindeki objeleri yaklaşık %25 oranında tespit edip sınıflandırabilmiştir. Şekil 11'de veri setinin eğitimi sonucunda elde edilen kayıp fonksiyonu grafiği gösterilmiştir.

YOLOv5 ve SSD_resnet101 modellerinin DOTA veri seti ile eğitim sonuçları değerlendirildiğinde ortalama doğruluk oranı olarak YOLOv5 modeli daha başarılı bulunmuştur. Ancak modellerin eğitim süreleri göz önüne alındığında SSD modelinin eğitimini daha kısa sürede tamamladığı görülmüştür. Modellerin DOTA veri seti üzerinde nesne tespitindeki başarı kıyaslaması Tablo 2'de gösterilmiştir.



Şekil 11. SSD_resnet101 modelinin 35000 adım eğitim sonrası kayıp değeri grafiği

Tablo 2. YOLOv5 ve SSD_resnet101 algoritmalarının DOTA veri seti üzerinde başarı kıyaslaması

Model Adı	Görüntü Boyutu (Image Size)	Yığın Sayısı (Batch Size)	Devir Sayısı (Epoch)	Eğitim Süresi	Ortalama Doğruluk Oranı (mAP)
YOLOv5 Pytorch	640x640	16	100	7 saat 15 dakika	0,567
SSD_resnet101_Tensorflow	640x640	4	100	4 saat 32 dakika	0,255

SSD modelinin eğitim süresinin daha kısa olmasına karşın ortalama doğruluk oranının oldukça düşük olduğu görülmüştür. Veri seti içerisindeki görüntüler üzerinde bulunan nesnelerin etiket sayıları ve modelin parametreleri incelenmiştir. Bu inceleme sonrasında veri setinin dengesiz (unbalanced) bir yapıya sahip olduğu anlaşılmıştır. Veri içerisinde bulunan dengesiz dağılımın eğitim sonucunu olumsuz etkileyebileceği değerlendirilmiştir. Ayrıca DOTA veri seti uydu sistemlerinden yani yerden yüksek mesafelerden elde edilen görüntülerden oluştuğu için görüntüler üzerinde bulunan nesneler çok küçük boyuttadır. Bu sebeple eğitim sonucunun istenilen seviyede olamadığı ikinci bir neden olarak değerlendirilmiştir.

Görüntüler üzerinde bulunan nesnelerin boyutlarını bir nebze arttırmak adına veri ayırma (image split) yöntemi DOTA veri setine uygulanmıştır. Daha sonra bazı sınıflar özelinde veri büyütme işlemi (data augmentation) uygulanarak veri seti içerisindeki dengesiz dağılım azaltılmıştır. Oluşturulan bu veri seti yaklaşık olarak 7.000 eğitim veri seti, 3.000 doğrulama veri seti şeklinde bölünmüştür. Yeni veri seti bu haliyle tekrardan modeller üzerinde eğitime tabii tutulmuştur. Modellerin yeni veri seti üzerindeki sonuçları Tablo 3'te gösterilmektedir.

Tablo 3. YOLOv5 ve SSD_resnet101 algoritmalarının yeni veri seti başarı kıyaslaması

Model Adı	Görüntü Boyutu (Image Size)	Yığın Sayısı (Batch Size)	Devir Sayısı (Epoch)	Eğitim Süresi	Ortalama Doğruluk Oranı (mAP)
YOLOv5 Pytorch	640x640	16	100	24 saat 20 dakika	0,867
SSD_resnet101_Tensorflow	640x640	4	100	7 saat 54 dakika	0,338

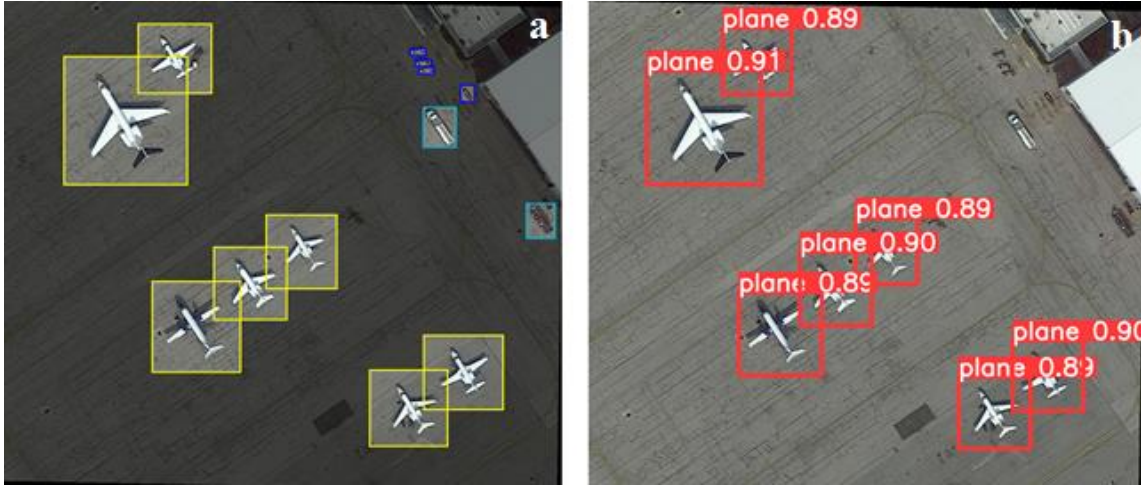
Tablo 3’ te gösterilen eğitim sonuçları göz önüne alındığında görüntü bölme ve veri büyütme uygulanan veri setinde YOLOv5 algoritmasının ortalama doğruluk oranı % 52.9 oranında artarak 0.867 değerine ulaşmıştır. Eğitim süresi veri setinin büyümesinden kaynaklı artarak 24 saat 20 dakika olarak kaydedilmiştir. SSD algoritmasında ise ortalama doğruluk oranı % 32.5 oranında artarak 0.338 değerine ulaşmıştır. Eğitim süresi YOLOv5 algoritmasına oranla daha az artış göstererek 7 saat 54 dakika sürmüştür. Veri setinde yapılan iyileştirmeler sonrasında her iki nesne tespit algoritmasının da ortalama doğruluk oranlarında ciddi artış gözlenmiştir.

Deneysel çalışmaların sonunda literatüre katkısı olması açısından DOTA veri seti giriş görüntüsü değiştirilerek YOLOv5 modelin başarımı gözlemlenmiştir. Başarım sonuçları Tablo 4’te verilmiştir. Sonuçlar göz önüne alındığında giriş görüntü boyutu arttıkça ortalama doğruluk oranının arttığı görülmüştür. Eğitim süresinin de artışı nedeniyle nesne tespiti açısından YOLOv5 algoritması seçiminde bu etkiye dikkat edilmelidir.

Tablo 4. YOLOv5 algoritmasının farklı boyutlarda giriş görüntüleri üzerinde etkisi

Model Adı	Görüntü Boyutu (Image Size)	Yığın Sayısı (Batch Size)	Devir Sayısı (Epoch)	Eğitim Süresi	Ortalama Doğruluk Oranı (mAP)
YOLOv5 Pytorch	416 x 416	16	100	2 saat 15 dakika	0,47163
YOLOv5 Pytorch	640 x 640	16	100	7 saat 12 dakika	0,56781
YOLOv5 Pytorch	720 x 720	12	100	10 saat 20 dakika	0,5764

Şekil 12’de test görüntüsü ile bu test görüntüsünün eğitilen YOLOv5 modeli ile tahmin görüntüsü verilmiştir. Test görüntüsünde bulunan “küçük araç” ve “büyük araç” sınıflarına ait nesnelerin tahminleri başarısız olmuştur. “Uçak” sınıfına ait nesnelerin tahmini yaklaşık %90 oranında başarılı olmuştur.



Şekil 12. Etiketli Test Görüntüsü (a) ve YOLOv5 Modelin Tahmin Çıktısı Görüntüsü (b)

5. Sonuç ve Öneriler

Bu çalışma ile tek aşamalı nesne tespit modellerinden YOLOv5 ve SSD algoritmaları kullanılarak uydu sistemleri vasıtasıyla havadan elde edilen görüntüler üzerinde bulunan nesnelerin tespit edilmesi amacıyla bir yazılım geliştirilmiştir. Çalışmada hedeflenen amaç doğrultusunda geliştirilen uygulamanın başarısı her iki algoritma performans ve verimlilik açısından analiz edilmiştir.

Nesne tespiti amacıyla birçok evrimsel sinir ağı modeli geliştirilmiştir. Her modelin farklı çalışma formatları bulunmaktadır. Bu nedenle kullanılacak nesne tespit modeline göre DOTA veri seti görüntü formatlarında değişiklikler yapılmıştır.

Literatüre bakıldığında birçok nesne tespit çalışmasının olduğu görülmüştür. Genellikle veri seti olarak MS COCO, PASCAL VOC ve ImageNet gibi günlük yaşamdan çok yönlü görüntüler içeren zengin veri setleri kullanılmaktadır. DOTA veri setinin bu veri setlerinden en temel ve en büyük farkı bir perspektiften oluşan görüntüler içermesidir. Her ne kadar YOLOv5 modelinde iyi denebilecek sonuçlar üretilse de çok iyi sonuçlar üretememesinin nedeni bu temel farktır.

Veri büyütme ve görüntü bölme yöntemlerinin uygulanması sonrasında her iki model içinde sonuçlar olumlu yönde olmuştur. İlk aşama göz önüne alındığında dengesiz dağılıma sahip bir veri seti üzerinde YOLOv5 algoritmasının bu olumsuz durumun üstesinden gelebildiği değerlendirilmiştir. Giriş görüntüsü boyutlarında artış sağlanarak eğitime tabi tutulan YOLOv5 algoritması için eğitim sonrasında ortalama doğruluk değerlerinin arttığı değerlendirilmiştir.

Performans ve verimlilik açısından her geçen gün artış gösteren nesne tespit algoritmalarının günümüzde öncelikli kullanım amacı İHA'lar olarak kabul edilmektedir. Bu sebeple bu modellerin başarılarının artırılması adına çalışmalar hız kazanmaktadır. Önümüzdeki dönemlerde verimi ve etkinliği göz önüne alındığında özellikle YOLOv5 modeli ile havadan elde edilen veriler kullanılarak daha kapsamlı uygulamalar yapılacaktır. Nesne tespit ve tanıma işleminin yanı sıra görüntüde hangi nesneden kaç tane olduğuna dair ek çalışma yapılarak kullanılan Silahlı İnsansız Hava Araçları özelinde "Tehdit Değerlendirme Silah Tahsis" algoritmalarına bu çalışma ile girdi sağlanması hedeflenmektedir.

Yazarların Katkıları

KT ve FH deneysel çalışmaları birlikte yapmıştır. KT çalışmayı dokümanete etmiştir. Yazarlar bu çalışmanın son halini okudu ve onayladı.

Çıkar Çatışması

Yazarlar, çıkar çatışması olmadığını beyan eder.

Kaynaklar

- [1]. Mihajlovic, I. (2021, September 24). Everything you ever wanted to know about computer vision. here's a look why it's so awesome. Medium. Acces Date May 19, 2022, from <https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e>
- [2]. R. Comert, U. Avdan and E. Senkal, "Unmanned aerial vehicles usage areas and future expectations," "İnsansız hava araçlarının kullanım alanları ve gelecekteki beklentiler," IV. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu, Zonguldak, 2012.
- [3]. Condliffe, J. (2015, January 9). This was the first ever satellite image of the entire U.S. Gizmodo. Acces Date May 19, 2022, from <https://gizmodo.com/this-was-the-first-ever-satellite-image-of-the-entire-u-1678463325>
- [4]. Wikimedia Foundation. (2022, May 13). Satellite imagery. Wikipedia. Acces Date May 19, 2022, from https://en.wikipedia.org/wiki/Satellite_imagery
- [5]. News bureau. News Bureau. (n.d.). Acces Date May 19, 2022, from <https://www.maxar.com/news-bureau>
- [6]. Wang, J., Guo, W., Pan, T., Yu, H., Duan, L., & Yang, W. "Bottle Detection in the Wild Using Low-Altitude Unmanned Aerial Vehicles." IEEE 21st International Conference on Information Fusion (FUSION 2018), 2018.
- [7]. Toraman S. "Derin Öğrenme ile İnsansız Hava Aracı Görüntülerinden Yaya Tespiti." Journal of Aviation, 2018, 2(2), 64-69.
- [8]. Kılıç, E. & Öztürk, S. "İnsansız Hava Aracı Görüntülerinde Evrişimli Sinir Ağları Tabanlı Araba Sayımı Yöntemi ." Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 2020, 36(2), 170-177.
- [9]. E. Bayhan, Z. Ozkan, M. Namdar and A. Basgumus, "Deep Learning Based Object Detection and Recognition of Unmanned Aerial Vehicles," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, 1-5
- [10]. H. S. Dıkbayır, H. İ. Bülbül . "Derin Öğrenme Yöntemleri Kullanarak Gerçek Zamanlı Araç Tespiti ." TÜBAV Bilim Dergisi, 2020, 13(3), 1-14 .
- [11]. U. Talaş, U. Yüzgeç, B. Çubukçu. "Derin Öğrenme ile Nesne Tanıyan Robot." Avrupa Bilim ve Teknoloji Dergisi, 2021, (31), 127-133.
- [12]. Şeker, A. , Diri, B. & Balık, H. H. "Derin Öğrenme Yöntemleri Ve Uygulamaları Hakkında Bir İnceleme" Gazi Mühendislik Bilimleri Dergisi ,2017, 3 (3) , 47-64 .
- [13]. Süslü, A. "Doğa ve İnsan Bilimlerinde Yapay Zekâ Uygulamaları." Akademia Doğa ve İnsan Bilimleri Dergisi, 2019, 5(1) , 1-10 .
- [14]. Cengil, E. & Çınar, A. "A New Approach For Image Classification: Convolutional Neural Network ." European Journal of Technique(EJT), 2016, 6(2), 96-103.
- [15]. G. S. Xia, X. Bai, J. Ding, Z. Zhu, S. Belongie, J. Luo, M. Datcu, M. Pelillo ve L. Zhang, "Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021,1.
- [16]. A. L. Morera, A. Sanchez, A.B Moreno, A. D. Sappa "SSD vs. YOLO for Detection of Outdoor Urban Advertising Panels under Multiple Variabilities.", Sensors, 2020 20. 4587. 10.3390/s20164587.
- [17]. T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. "Focal loss for dense object detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(2), 318–327.